Anatomy of energetic changes accompanying urea-induced protein denaturation

Matthew Auton, Luis Marcelo F. Holthauzen, and D. Wayne Bolen[†]

Department of Biochemistry and Molecular Biology, University of Texas Medical Branch, 301 University Boulevard, 5.154 Medical Research Building, Galveston, TX 77555-1052

Because of its protein-denaturing ability, urea has played a pivotal role in the experimental and conceptual understanding of protein folding and unfolding. The measure of urea's ability to force a protein to unfold is given by the *m* value, an experimental quantity giving the free energy change for unfolding per molar urea. With the aid of Tanford's transfer model [Tanford C (1964) *J Am Chem Soc* 86:2050–2059], we use newly obtained group transfer free energies (GTFEs) of protein side-chain and backbone units from water to 1 M urea to account for the *m* value of urea, and the method reveals the anatomy of protein denaturation in terms of residue-level free energy contributions of groups newly exposed on denaturation. The GTFEs were obtained by accounting for solubility and activity coefficient ratios accompanying the transfer of glycine from water to 1 M urea. Contrary to the opinions of some researchers, the GTFEs show that urea does not denature proteins through favorable interactions with nonpolar side chains; what drives urea-induced protein unfolding is the large favorable interaction of urea with the peptide backbone. Although the *m* value is said to be proportional to surface area newly exposed on denaturation, only ≈25% of the area favorably contributes to unfolding (because of newly exposed backbone units), with ≈75% modestly opposing urea-induced denaturation (originating from side-chain exposure). Use of the transfer model and newly determined GTFEs achieves the long-sought goal of predicting urea-dependent cooperative protein unfolding energetics at the level of individual amino acid residues.

m value | transfer free energy | transfer model | activity coefficient | self-avoiding random coil

Understanding the energetics of protein–solute interactions is one of the most elusive goals of protein science, with urea-induced denaturation serving as a long-standing reminder of the inability to experimentally account for such fundamental interactions in a detailed manner. More than 40 years ago, Tanford set out to identify the sites and free energies of urea interaction with protein groups in enough detail to account for the energetics of urea-induced denaturation (1–3). In principle, the transfer model he developed provides a means of dissecting which groups (side chain and backbone) are involved in denaturation and how much they contribute to the free energy of denaturation by urea. For the transfer model (see Scheme 1) to be successful, two requirements must be met: (*i*) accurate transfer free energy changes for side-chain and backbone groups must be known, and (*ii*) the free energy of transfer of a native or denatured state of a protein from water to the urea solution must be equal to the sum of the transfer free energy contributions of its solvent-exposed parts. Of these two requirements, the ability to obtain accurate transfer free energies of side chains and backbone groups has been a particularly difficult impediment to quantifying the energetics of urea-induced denaturation. Here, we present results for the transfer free energy of glycine from water to 1 M urea that improve the evaluation of side-chain transfer free energies, and elevate the transfer model to the quantitative level originally sought by Tanford and colleagues.

The standard state transfer free energy of an amino acid (*i*) from water to urea ($\Delta G^\circ_{\mathrm{tr},i}$) is given by Eq. **1**, where $a_{i,\mathrm{w}}$ and $a_{i,\mathrm{urea}}$ represent the thermodynamic activities of amino acid *i* at its solubility limit in water and urea, respectively (1). Terms ($\gamma_{i,\mathrm{w}}/\gamma_{i,\mathrm{urea}}$) and ($s_{i,\mathrm{w}}/s_{i,\mathrm{urea}}$) represent the activity coefficient and concentration ratios in either mole fraction, molar, or molal units, each quantity being evaluated at the solubility limit of amino acid *i*.

$$\Delta G^\circ_{\mathrm{tr},i} = RT \ln(a_{i,\mathrm{w}}/a_{i,\mathrm{urea}})$$
$$= RT \ln(s_{i,\mathrm{w}}/s_{i,\mathrm{urea}}) + RT \ln(\gamma_{i,\mathrm{w}}/\gamma_{i,\mathrm{urea}}). \quad \textbf{[1]}$$

Historically, that part of the transfer free energy arising from concentration ratios was taken to represent the transfer free energy, with the contribution due to the activity coefficient ratio considered as small or negligible (1). Because the data on amino acid solubility in water and urea are reported without activity coefficients, the transfer free energies of amino acids presented previously must be regarded as *apparent* transfer free energies [$\Delta G^{\mathrm{app}}_{\mathrm{tr},i} = RT \ln(s_{i,\mathrm{w}}/s_{i,\mathrm{urea}})$] (1, 4).

The transfer free energies of side-chain groups from water to 1 M urea are needed in application of the transfer model. Because the structure of glycine is a common part of the structure of all amino acids, Nozaki and Tanford suggested the side-chain or group transfer free energies (GTFEs) could be obtained by subtracting the apparent transfer free energy of glycine ($\Delta G^{\mathrm{app}}_{\mathrm{tr,Gly}}$) from the apparent transfer free energy of each amino acid *i* (1). Thus, the GTFEs reported by Nozaki and Tanford are also *apparent* side-chain transfer free energy changes, $\mathrm{GTFE}^{\mathrm{app}}_{\mathrm{tr},i}$ (1). Here we show that use of the original $\mathrm{GTFE}^{\mathrm{app}}_{\mathrm{tr},i}$ values with the transfer model does not result in quantitative results or predictive capability that would encourage use of the model. However, activity coefficient data are available for glycine in water and 1 M urea and can be combined with the glycine solubility ratio to give $\Delta G^\circ_{\mathrm{tr,Gly}}$. Subtraction of $\Delta G^\circ_{\mathrm{tr,Gly}}$ from $\Delta G^{\mathrm{app}}_{\mathrm{tr},i}$ values of the amino acids results in side-chain group transfer quantities we call $\mathrm{GTFE}^*_{\mathrm{tr},i}$. These quantities account for the activity coefficients of glycine and their use makes the transfer model highly predictive of thermodynamic quantities derived from urea-induced denaturation of proteins. This revitalization of the transfer model provides insight into the anatomy of energetic contributions of side chains and backbone to the cooperativity (*m* value) of urea-induced protein unfolding, plac-

ing the model on a firm foundation for further predictions of the thermodynamic properties of urea–protein interactions.

Considerable diversity of opinion exists concerning mechanisms of urea-induced denaturation of proteins. Some authors cite the favorable interaction of urea with nonpolar groups as the source of its denaturing action (1, 3, 5, 6), others claim that urea's favorable interaction with peptide backbone is responsible (7–9), and still others refer to its favorable interactions with both the backbone and nonpolar groups as the cause of denaturation (10–12). Given the major role urea has played in understanding the thermodynamics and kinetics of protein folding/unfolding, the inability to achieve consensus on how it denatures proteins is symptomatic of the limited and sometimes misleading experimental model compound data on urea interactions with protein groups. The long-standing conclusion that urea denatures proteins because of its favorable interaction with nonpolar side chains arises from the $\text{GTFE}_{\text{tr},i}^{\text{app}}$ values of nonpolar groups reported by Nozaki and Tanford (1). When activity coefficients are used to correctly account for the transfer free energy of glycine, the interaction free energies between urea and nonpolar groups are not of a magnitude that justifies that long-standing conclusion. Our analysis using $\text{GTFE}_{\text{tr},i}^{*}$ values and the transfer model clearly shows that urea's favorable interaction with peptide backbone is the driving force for urea-induced denaturation, with nonpolar group–urea interactions playing little or no role in the process.

Results

Group Transfer Free Energies. A special case of Eq. **1** applies if the solubility of amino acid i reaches a low enough limit in water and in urea solution that the activity coefficient of the amino acid takes a value of unity. Under this condition, the chemical activity of the amino acid becomes equal to its concentration and $\Delta G_{\text{tr},i}^{\text{app}} = \Delta G_{\text{tr},i}^{\circ}$. Activity coefficients are difficult to measure, so to obviate the need for measurement, model compounds may be chosen for transfer free energy measurements that have solubility limits low enough that activity coefficients can be ignored. Evidence in support of reaching such limits occurs when the apparent transfer free energy of the chemical moiety of interest is found not to depend on the particular (highly insoluble) model compound containing the moiety. Such chemical-model-independent transfer free energies were established for the peptide backbone moiety, not only from water to 1 M urea, but also from water to an array of protecting osmolytes at 1 M concentration (13).

Traditionally, the compounds used as models of side-chain moieties have been the amino acids themselves, and not all of these meet the criterion of having low solubility limits. Fig. 1 provides the molar and molal water solubility limits of 19 amino acids ranked from the most (proline) to the least (tyrosine) soluble. As a point of reference, cyclic glycylglycine is a model compound for the peptide backbone moiety whose solubility limit is low enough that $(\gamma_{i,\text{w}}/\gamma_{i,\text{urea}}) \rightarrow 1$. Given that molar and molal concentrations become identical in dilute solution, the difference between these two concentration scales provides a visual indication of the relative solubility limits of the amino acids as they become dilute enough for the two scales to coincide. Coincidence of molar and molal scales does not necessarily mean the compound meets the solubility limit where the activity coefficient is unity; it is used here merely as a visual indicator of the approach to dilute solution.

Although the solubility limits of amino acids to the right of threonine in Fig. 1 are indistinguishable in terms of their molar and molal concentrations, glycine, alanine, and amino acids to their left are sufficiently soluble to exhibit molal and molar concentration differences. For amino acids to the left of threonine, one might expect molal and molar activity coefficients unequal to one. Rafflenbeul *et al.* (14) have measured molal
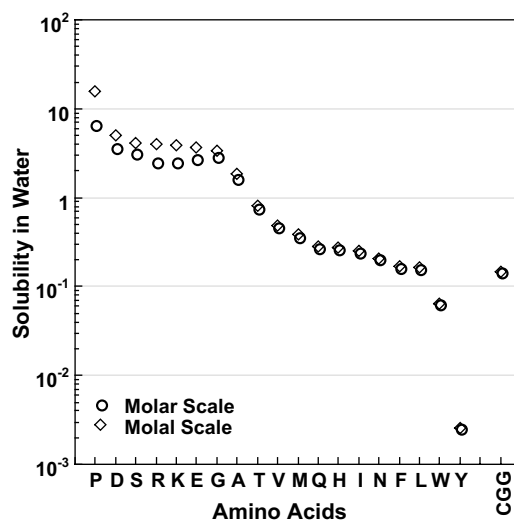

Fig. 1. The solubility of amino acids in water in the molar (○) and molal (◇) concentration scales are ranked from greatest (proline) to least (tyrosine) according to molal solubility. The solubility of the peptide backbone model compound, cyclic glycylglycine (CGG), is included on the far right.

activity coefficients for glycine in urea and water, from which the authors obtained molar activity coefficients of glycine at its solubility limit in water (0.844) and in 1 M urea solution (0.904). This gives an activity coefficient ratio of 0.934 and a contribution of -40.3 cal·mol^{-1}·M^{-1} to the transfer free energy of glycine, due to the activity coefficient ratio term (see Eq. **1**). Note that the activity coefficient ratio is clearly closer to unity than the individual activity coefficients. In general, the activity coefficient ratio will converge to unity significantly before the individual activity coefficients reach unity.

As mentioned previously, the procedure to obtain the group transfer free energies of individual amino acid side chains is to subtract the transfer free energy of glycine from the transfer free energy of all other amino acids (see Eq. **2**), where $\text{GTFE}_{\text{tr},i}^{\circ,*,\text{app}}$ is the transfer free energy of the side chain of amino acid i (1). $\text{GTFE}_{\text{tr},i}^{\circ,*,\text{app}}$ becomes $\text{GTFE}_{\text{tr},i}^{\text{app}}$ if only terms 1 and 3 on the right side of Eq. **2** are considered important, $\text{GTFE}_{\text{tr},i}^{*}$ if only terms 1,3, and 4 are important, and $\text{GTFE}_{\text{tr},i}^{\circ}$ if all four terms are important. The first two terms on the right side of Eq. **2** comprise the standard state transfer free energy of amino acid i ($\Delta G_{\text{tr},i}^{\circ}$), and the two terms within the brackets represent the standard state glycine transfer free energy ($\Delta G_{\text{tr,Gly}}^{\circ}$) (1).

$$\text{GTFE}_{\text{tr},i}^{\circ,*,\text{app}} = RT \ln\left(\frac{s_{i,\text{w}}}{s_{i,\text{urea}}}\right) + RT \ln\left(\frac{\gamma_{i,\text{w}}}{\gamma_{i,\text{urea}}}\right)$$
$$- \left[RT \ln\left(\frac{s_{\text{Gly,w}}}{s_{\text{Gly,urea}}}\right) + RT \ln\left(\frac{\gamma_{\text{Gly,w}}}{\gamma_{\text{Gly,urea}}}\right) \right].$$

$$[2]$$

Because activity coefficient data for glycine in water and urea are available in the literature, we take the opportunity to report the contributions of both terms in the brackets of Eq. **2**, which together give the standard state glycine transfer free energy, $\Delta G_{\text{tr,Gly}}^{\circ}$. Combining $\Delta G_{\text{tr,Gly}}^{\text{app}} = 18.7$ cal/mol reported from the glycine solubility data of Nozaki and Tanford (1), with the activity coefficient ratio contribution determined by Rafflenbeul *et al.* (14) (-40.3 cal·mol^{-1}·M^{-1}), gives $\Delta G_{\text{tr,Gly}}^{\circ} = -21.7$ cal·mol^{-1}·M^{-1}. Thus, we subtract -21.7, whereas Nozaki and Tanford subtracted $+18.7$ cal·mol^{-1}·M^{-1} from each of the side-chain-containing amino acids. Including the glycine activity
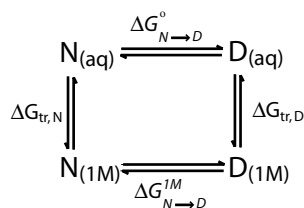
Scheme 1. The transfer model

coefficient ratio term changes the side-chain GTFEs for water to urea transfer considerably from what has been used and cited since the 1960s.

Transfer Model. The transfer model, shown in Scheme 1, compares the extent to which the native (N) and denatured (D) state equilibrium in 1 M urea (as given by $\Delta G_{N\to D}^{1M}$) differs from $N_{(aq)} \rightleftarrows D_{(aq)}$ in buffer solution (given by $\Delta G_{N\to D}^{\circ}$). The difference in free energy for the two reactions, $\Delta G_{N\to D}^{1M} - \Delta G_{N\to D}^{\circ}$, represents the *m* value, a quantity obtained experimentally that measures the efficacy of urea in denaturing a protein. The thermodynamic cycle results in the relationship $\Delta G_{N\to D}^{1M} - \Delta G_{N\to D}^{\circ} = \Delta G_{tr,D} - \Delta G_{tr,N}$, showing that the free energy difference of the perpendicular reactions, $\Delta G_{tr,D} - \Delta G_{tr,N}$, is also equal to the *m* value (15). A goal of the transfer model is to evaluate $\Delta G_{tr,N}$ and $\Delta G_{tr,D}$ individually and from their difference predict the experimental *m* value of the protein.

$\Delta G_{tr,D}$ and $\Delta G_{tr,N}$ represent the transfer free energies of denatured and native protein from buffer solution to 1 M urea, and with knowledge of GTFEs of side chains and peptide backbone, N and D transfer free energies can be evaluated as shown previously (15). Beginning with Protein Data Bank structural coordinates of the N state of a particular protein, the solvent-exposed surface of the protein native state is determined with a probe of radius 1.4 Å. The accessible surface area for each type of backbone and side-chain group is summed and normalized against the area of the group exposed to solvent in a Gly-Xaa-Gly peptide. This gives the number of such groups on the protein exposed to solvent (15). The numbers of each type of solvent-exposed group are multiplied by the corresponding GTFE for that particular type of group, and all contributions are summed to give $\Delta G_{tr,N}$.

Likewise for the D state, the evaluation of $\Delta G_{tr,D}$ requires additivity of component groups as well as a model of the denatured state solvent accessibility. Creamer *et al.* (16, 17) considered models for two denatured state extremes, a random coil in a good solvent, representing a highly solvent-exposed expanded state, and a compact denatured state, representing a denatured state with a high degree of solvent inaccessibility. For these extreme models, the solvent accessibilities of groups and the products of their GTFEs and numbers exposed were used to calculate $\Delta G_{tr,D}$ for the two extreme models. Schellman's (18) model for the denatured ensemble is half-way between the two extremes, and we have adopted that model for use in this work.

Predictions. The effect of using the historical GTFE$_{tr,i}^{app}$ values to predict *m* values of urea-induced denaturation of monomeric proteins can be seen in Fig. 2*A*, with molecular weights of proteins increasing roughly from right to left in the figure. It is observed that the predicted *m* values of urea-induced denaturation do not at all agree with experimentally obtained *m* values, and the prediction becomes markedly worse as the size of the protein increases. For comparison, filled circles represent disulfide-containing proteins. By contrast, Fig. 2*B* shows the predicted *m* values for urea denaturation of the same proteins based on GTFE$_{tr,i}^{*}$ values. The one-to-one correspondence be-
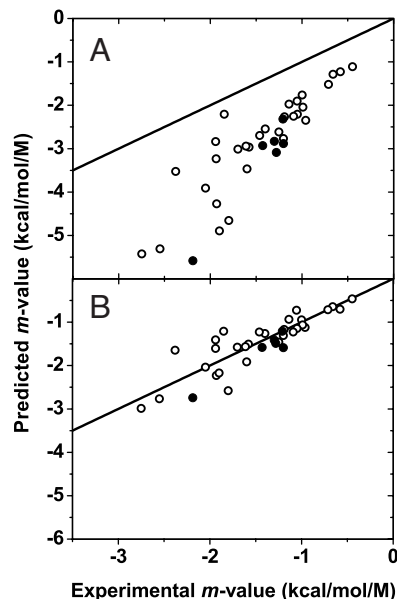

Fig. 2. Predicted versus experimentally determined *m* values. The proteins represented are listed in the *SI Appendix*. Disulfide-containing proteins are represented as filled circles. The identity line is shown for comparison. Predicted *m* values were obtained by using GTFE$_{tr,i}^{app}$ values (*A*) and the corrected GTFE$_{tr,i}^{*}$ values (*B*) as explained in the text.

tween predicted and observed *m* values shows considerable improvement in prediction as a result of using GTFEs based on $\Delta G_{tr,Gly}^{\circ}$ instead of $\Delta G_{tr,Gly}^{app}$.

We consider that an important source of deviations from the identity line in Fig. 2*B* may be differences between the degree of residual structure in the actual denatured state ensembles of proteins displayed, in comparison with the model we used to represent the denatured state. To compare denatured state models, *m* values were predicted assuming that the D state is a random coil in a good solvent (upper-bound model) or a compact denatured state (lower-bound model). The results of linear fits of the two extreme models are shown in Fig. 3. Schellman's model for the denatured state, used here, results in one-to-one correspondence between predicted and observed *m* values and is clearly superior to the upper- and lower-bound denatured state models in predicting urea-induced protein denaturation *m* values.
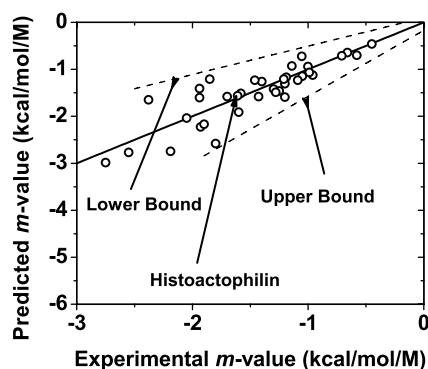

Fig. 3. Predicted *m* values based on GTFE$_{tr,i}^{*}$ values versus experimentally determined *m* values. The dashed lines indicated give linear fits of the lower- and upper-bound calculated *m* values. The solid line is the identity line for predicted and experimental *m* values. The position of histoactophilin, the protein example used in Fig. 5, is shown for reference.
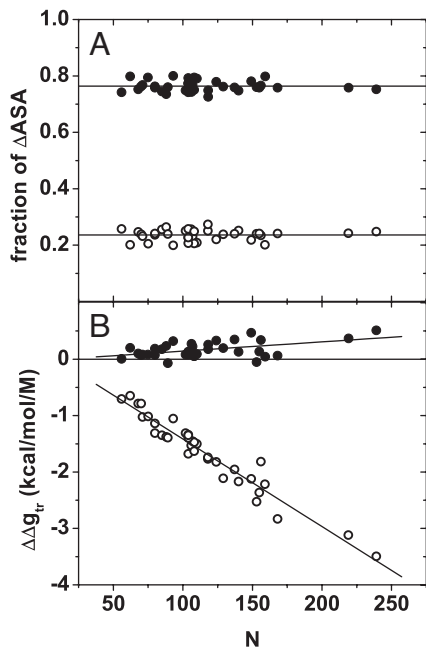
Fig. 4. Fractional surface areas and *m* value contributions of groups newly exposed on unfolding. (*A*) Fractions of ΔASA, in which ASA is solvent-accessible surface area, contributed by groups newly exposed on unfolding are shown as a function of protein size (number of residues). The lines represent the average value for side chains and backbones. (*B*) *m* value contributions of side chains and backbone units (ΔΔg_{tr}) versus number of residues. The slopes of the fitted lines are $(1.6 \pm 0.5) \times 10^{-3}$ for the side chains and $(-15.5 \pm 0.7) \times 10^{-3}$ for the backbones. For both *A* and *B*, the values for the side chains are shown as filled circles and for the backbone as open circles.

On protein unfolding, side-chain and peptide backbone units become newly exposed and it is the sum of transfer free energies from water to 1 M urea of these groups that gives the predicted *m* value. Fig. 4*A* shows that ≈75% of the newly exposed surface area on protein unfolding originates from side chains with ≈25% coming from peptide backbone. These percentages are rather constant with protein chain length. It may come as a surprise to learn that the groups that contribute most to newly exposed surface area do not contribute favorably to denaturation by urea. Fig. 4*B* shows that the favorable interaction of urea with peptide backbone is responsible for denaturation and that urea interaction with newly exposed side chains, in fact, opposes protein unfolding to a small degree.

The results shown in Fig. 4*B* stand in contrast to the position held by many experimentalists for >40 years, that urea denatures proteins because of the *favorable* interaction between urea and nonpolar side chains exposed on unfolding. That long-held view was based on GTFE$_{tr,i}^{app}$ values that overestimate the urea–side-chain transfer free energies by 40 cal·mol^{-1}·M^{-1}. With GTFE$_{tr,i}^{*}$ values the nonpolar side chains are relatively unimportant factors in urea-induced denaturation. The supporting information (SI) *Appendix* provides a table of GTFE$_{tr,i}^{*}$ and GTFE$_{tr,i}^{app}$ values for side chains and backbone, a listing of monomeric proteins with experimental *m* values used in the figures, and evaluation of the activity coefficients for glycine in water and 1 M urea on which the GTFE$_{tr,i}^{*}$ results are based.

Because the transfer model gives detailed information on the groups that become newly exposed on urea-induced unfolding, it is possible to reveal the sources and amounts of free energy arising from solvent exposure of individual parts of the protein undergoing denaturation. Fig. 5 displays the free energy of groups newly exposed on urea-induced denaturation of histoactophilin. The abscissa gives the surface areas of the groups that
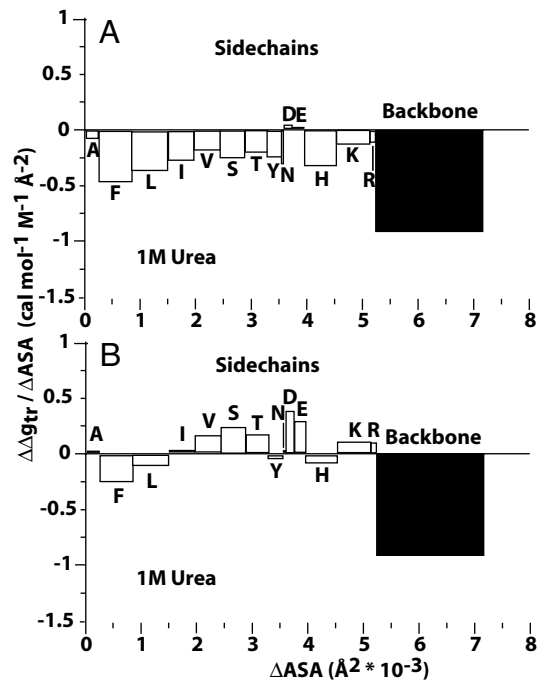

Fig. 5. Side-chain and backbone *m* value contributions (ΔΔg_{tr}) divided by their respective contribution to ΔASA (cal·mol^{-1}·M^{-1}·Å^{-2}) as a function of total surface area newly exposed on denaturation of histoactophilin. Open bars are for side chains, and black bars are for peptide backbone units. The contributions were calculated by using GTFE$_{tr,i}^{app}$ (*A*) and GTFE$_{tr,i}^{*}$ (*B*).

become newly exposed on urea-induced denaturation. The ordinate is the free energy contribution per square ångström of the individual protein groups, with the corresponding areas representing the free energy contribution of that type group to the overall predicted *m* value. The algebraic sum of all free energy contributions in the panel equals the predicted *m* value for the protein.

Discussion

Transfer free energy measurements are sources of the concepts that urea denatures proteins through its favorable interactions with nonpolar groups and through its favorable interaction with peptide backbone (1, 3, 19). Recently, other techniques have been used to assess urea's interactions with protein groups in an attempt to account for its denaturing action. Molecular dynamics simulations of proteins in urea indicate that urea interacts with exposed polar groups by means of hydrogen bonding (20–22). The potential of mean force approach of O'Brien *et al.* (8) suggested that hydrophobic association is not significantly changed by urea and that urea's denaturing action is due to its direct interaction with the peptide backbone. And by using preferential interaction and parameterization methods, Record and coworkers (7, 23, 24) concluded that dominant interaction of urea is with the polar amide surface of proteins, and interactions with nonpolar and anionic side chains are an order of magnitude smaller. The results described here are in substantial agreement with these more recent conclusions.

In principle, the transfer model proposed by Tanford enables identification and dissection of the free energy contributions of those parts of the protein participating in the denaturant-induced transition. Such an understanding of protein–solvent interactions has been a goal in protein science since the 1930s (25). The model provides a framework for understanding denaturants and predicting protein stability, but only if several assumptions and conditions are met. Chief among these assump-

Auton *et al.*

tions is that the free energy of protein transfer from water to denaturing solvent consists of a simple sum of the transfer free energies of solvent-exposed parts of the protein. Of course, even if additivity does hold, the model will only be as successful as the accuracy of the side-chain and backbone transfer free energies used, and the ability to evaluate solvent exposure of the native and denatured species. The inability of Tanford to obtain adequate side-chain and backbone transfer free energy values from water to urea solution and the lack of tools to evaluate solvent exposure of native and denatured protein resulted in protein stability predictions that failed at the time to compete with the success of the less enlightening denaturant binding and linear extrapolation models (3, 26).

Subsequent development of computational methods for evaluating the solvent exposure of native and denatured protein (16, 17, 27), and for evaluating the peptide backbone transfer free energies that exhibit additivity and are independent of the chemical model (13), improved chances for successfully applying the transfer model. Still, as seen in Fig. 2A, the model failed in prediction of the m values for urea-induced denaturation of proteins. Better evaluation of side-chain transfer free energies was lacking. In theory, the transfer free energies of the side chains can be obtained by subtracting the transfer free energy of glycine from the transfer free energy of each side-chain-containing amino acid (1). In practice, in the past, the activity coefficient ratio of glycine in 1 M urea and water was not accounted for, which resulted in reported side-chain transfer free energies some 40 cal·mol^{-1}·M^{-1} more favorable than they actually are. Use of the corrected side-chain transfer free energies (GTFE$^*_{\mathrm{tr},i}$) in applying the transfer model results in one-to-one correspondence between predicted and experimental m values for a wide variety of monmeric proteins reversibly denatured by urea. This result, shown in Fig. 2B, provides strong evidence that the newly determined side-chain and backbone GTFEs are markedly improved over those used in the past, and that group additivity holds to a significant degree.

Although GTFE$^*_{\mathrm{tr},i}$ values correctly account for the free energy of glycine transfer, the transfer free energies of the side-chain-containing amino acids remain *apparent* because of the unavailability of activity coefficients for side-chain-containing amino acids in water and urea. Why then should the GTFE$^*_{\mathrm{tr},i}$ values result in as good agreement between predicted and experimental m values as shown? It is clear from the transfer model that the m value depends on contributions only from groups buried in the native state that become newly exposed on denaturation. Such groups are mostly hydrophobic, with solubilities significantly lower than that of glycine. As the solubility limit of a compound decreases, its activity coefficient converges toward unity, and the activity coefficient ratios converge to unity well before the individual activity coefficients do (28). Thus, it appears that the activity ratio terms (Eq. **2**) for such residues tend to zero, and that GTFE$^*_{\mathrm{tr},i}$ values adequately define the side-chain group transfer free energies of the newly exposed side chains. Still, improvements in the transfer model are possible with knowledge of amino acid activity coefficients in water and 1 M urea.

The proteins used in this study were drawn from those reported in studies by Myers *et al.* and Hong *et al.*, along with some we added (7, 29–35). The m values depend on salt concentration and pH, so only those proteins are included that are monomeric, denature reversibly in the presence of urea at neutral pH and moderate salt, exhibit two-state behavior, and contain no cofactor or metal ion. It is evident that m values of some of the proteins deviate from the identity line in Fig. 2B. This deviation may be due to a variety of factors including deviation from two-state behavior, some degree of nonadditivity, inaccuracies in GTFE$^*_{\mathrm{tr},i}$ values, inaccurate experimental m values, unaccounted m value dependence on salt, and a urea-

denatured state that differs from the model used here. Although many potential factors could cause deviations or failure in agreement between predicted and experimental m values, the strong one-to-one agreement observed in Fig. 2B suggests that deviations from these sources are modest and random.

The denatured state model we have used (originally suggested by Schellman) has solvent-exposed surface area half-way between the upper- and lower-bound models of Creamer *et al.* (16–18). In Fig. 3, comparison of the results from use of the upper- and lower-bound models with results of Schellman's denatured state model shows that Schellman's model coincides so well with the identity line over the full set of data that it may be considered a consensus model for urea denaturation. What is lacking, however, is a dimensional and physical description of Schellman's denatured state model. In terms of surface area, Goldenberg (36) has recently shown that his self-avoiding random-coil denatured state model is itself half-way between the upper- and lower-bound models of Creamer *et al.* (16, 17), suggesting the physical description of a self-avoiding random coil for the thermodynamic denatured state model of Schellman (18).

Fig. 5 provides a graphical description of the predicted m value of histoactophilin, dissected into individual m value contributions of specific side-chain and backbone groups newly exposed on unfolding. Fig. 5A shows the individual m value contributions of newly exposed groups by using GTFE$^{\mathrm{app}}_{\mathrm{tr},i}$ values, quantities uncorrected for the glycine activity coefficient ratio. Many of these quantities were reported in 1963 and later (1, 4). Fig. 5A provides such information for urea-induced denaturation of histoactophilin by using GTFE$^{\mathrm{app}}_{\mathrm{tr},i}$ values. It is seen that newly exposed side chains and peptide backbone favorably interact with urea, giving the interpretation that both side chains and backbone are responsible for the ability of urea to denature proteins. When GTFE$^*_{\mathrm{tr},i}$ values are used instead, the results (see Fig. 5B) change markedly and provide no evidence of collective favorable interaction between nonpolar groups and urea. Although not all nonpolar side chains are represented in histoactophilin, proteins with an array of buried nonpolar side chains show little or no net m value contributions from nonpolar groups. It is clear from the algebraic sum of side-chain contributions that, overall, side chains contribute unfavorably to denaturation and the favorable urea–backbone interaction exclusively drives urea-induced denaturation.

The GTFE$^*_{\mathrm{tr},i}$ values markedly change interpretation and understanding of urea-induced denaturation of proteins in two ways. First, the reevaluated side-chain GTFEs no longer exhibit the highly favorable interactions between urea and nonpolar groups as concluded in the 1960s. Second, the dissection of the newly exposed groups on denaturation and their energy contributions to the m value provide detailed information on the origin of cooperative urea-induced unfolding of proteins. The groundbreaking work of Myers *et al.* (34) shows that experimental m values are proportional to the change in surface area on urea-induced protein unfolding. The transfer model puts a finer point on the observation by showing that newly exposed surface areas are not at all equal in the free energy they contribute toward protein unfolding. The peptide backbone contributes only ≈25% to the change in surface area on protein unfolding, but in terms of free energy it is entirely responsible for driving urea-induced unfolding of a protein. Importantly, newly exposed side chains contribute ≈75% of change in surface area on protein unfolding, but actually oppose denaturation by urea to a modest degree.

The original premise of Cohn, Edsall, and Tanford was that by knowing the transfer free energies of the individual parts of a protein and summing these contributions, one could predict thermodynamic behavior of proteins in the presence of different solvents/solutes (2, 25, 37). The result presented here shows that their original idea holds to a high degree.

BIOCHEMISTRY

Materials and Methods

Methods for determining the solubility ratios of amino acids and model compounds are well established, with the results of these ratios reported as transfer free energies from water to urea solutions (1, 4, 28, 38). As mentioned previously, because activity coefficient contributions to the transfer free energy of amino acid or model compounds were considered to be small or insignificant and were ignored (1), the transfer free energies based on solubility ratios alone must be considered *apparent* transfer free energies. Historically, side-chain GTFEs were obtained by subtraction of the *apparent* transfer free energy of glycine from that of each of the remaining amino acids to obtain the water-to-urea *apparent* GTFEs of each side chain, $GTFE_{tr,i}^{app}$ values.

Here, we use activity coefficient results from Rafflenbeul *et al.* (14) for glycine transfer from water to 1 M urea to obtain the standard state transfer free energy for glycine, $\Delta G_{tr,Gly}^{\circ}$. Extraction of molar activity coefficients of glycine in water and 1 M urea requires use of multiple terms and fitting coefficients of a polynomial expression, the details of which are given in *SI Appendix*. Subtraction of $\Delta G_{tr,Gly}^{\circ}$ from the apparent transfer free energies of the remaining amino acids gives a new set of GTFEs, $GTFE_{tr,i}^{*}$ values. In conjunction with the transfer model, $GTFE_{tr,i}^{*}$ and $GTFE_{tr,i}^{app}$ values reported in *SI Appendix* were used to predict the m values for urea-induced denaturation of proteins presented in this report. Details of the predictions are given in the *SI Appendix*.

Rafflenbeul *et al.* (14) also provide data for extraction of the standard state transfer free energy of alanine from water to 1 M urea, $\Delta G_{tr,Ala}^{\circ}$, by using the methods illustrated in the *SI Appendix*. The standard state GTFE for the methyl group, obtained by the subtraction, $\Delta G_{tr,Ala}^{\circ} - \Delta G_{tr,Gly}^{\circ}$, is presented in the $GTFE_{tr,i}^{*}$ table in *SI Appendix*.

1. Nozaki Y, Tanford C (1963) *J Biol Chem* 238:4074–4081.
2. Tanford C (1964) *J Am Chem Soc* 86:2050–2059.
3. Tanford C (1970) *Adv Protein Chem* 24:1–95.
4. Wang A, Bolen DW (1997) *Biochemistry* 36:9101–9108.
5. Kauzmann W (1959) *Adv Protein Chem* 14:1–63.
6. Wetlaufer DB, Malik SK, Stoller L, Coffin RL (1964) *J Am Chem Soc* 86:508–514.
7. Hong J, Capp MW, Saecker RM, Record MT, Jr (2005) *Biochemistry* 44:16896–16911.
8. O'Brien EP, Dima RI, Brooks B, Thirumalai D (2007) *J Am Chem Soc* 129:7346–7353.
9. Robinson DR, Jencks WP (1965) *J Am Chem Soc* 87:2470–2479.
10. Abeles RH, Frey PA, Jencks WP (1992) *Biochemistry* (Jones and Bartlett, Boston).
11. Creighton TE (1993) *Proteins: Structures and Molecular Properties* (Freeman, New York).
12. Fersht A (1998) *Structure and Mechanism in Protein Science* (Freeman, New York).
13. Auton M, Bolen DW (2004) *Biochemistry* 43:1329–1342.
14. Rafflenbeul L, Pang W-M, Schönert H, Haberle K (1973) *Z Naturforsch* 28:533–554.
15. Auton M, Bolen DW (2005) *Proc Natl Acad Sci USA* 102:15065–15068.
16. Creamer TP, Srinivasan R, Rose GD (1995) *Biochemistry* 34:16245–16250.
17. Creamer TP, Srinivasan R, Rose GD (1997) *Biochemistry* 36:2832–2835.
18. Schellman JA (2003) *Biophys J* 85:108–125.
19. Robinson DR, Jencks WP (1965) *J Am Chem Soc* 87:2462–2470.
20. Bennion BJ, Daggett V (2003) *Proc Natl Acad Sci USA* 100:5142–5147.
21. Caflisch A, Karplus M (1999) *Structure (London)* 7:477–488.
22. Tirado-Rives J, Orozco M, Jorgensen WL (1997) *Biochemistry* 36:7313–7329.
23. Courtenay ES, Capp MW, Record MT, Jr (2001) *Protein Sci* 10:2485–2497.
24. Courtenay ES, Capp MW, Saecker RM, Record MT, Jr (2000) *Proteins Suppl* 4:72–85.
25. Cohn EJ, McMeekin TL, Edsall JT, Weare JH (1934) *J Am Chem Soc* 56:2270–2282.
26. Pace CN (1975) *CRC Crit Rev Biochem* 3:1–43.
27. Lee B, Richards FM (1971) *J Mol Biol* 55:379–400.
28. Auton M, Bolen DW, *Methods Enzymol,* in press.
29. Auton M, Cruz MA, Moake J (2007) *J Mol Biol* 366:986–1000.
30. Baskakov IV, Bolen DW (1999) *Protein Sci* 8:1314–1319.
31. Holthauzen LM, Bolen DW (2007) *Protein Sci* 16:293–298.
32. Langdon GM, Jimenez MA, Genzor CG, Maldonado S, Sancho J, Rico M (2001) *Proteins* 43:476–488.
33. Mello CC, Barrick D (2003) *Protein Sci* 12:1522–1529.
34. Myers JK, Pace CN, Scholtz JM (1995) *Protein Sci* 4:2138–2148.
35. Wildes D, Anderson LM, Sabogal A, Marqusee S (2006) *Protein Sci* 15:1769–1779.
36. Goldenberg DP (2003) *J Mol Biol* 326:1615–1633.
37. Cohn EJ, Edsall JT (1943) *Proteins, Amino Acids, and Peptides as Ions and Dipolar Ions* (Reinhold, New York).
38. Bolen DW (2001) in *Methods in Molecular Biology*, ed Murphy KP (Humana, Totowa, NJ), pp 17–36.